Summer 8-1-1984

Continuity of the Null Space Basis and Constrained Optimization ; CU-CS-272-84

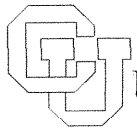
Richard H. Byrd
University of Colorado Boulder

Robert B. Schnabel
University of Colorado Boulder

Continuity of the Null Space Basis and Constrained Optimization *

Richard H. Byrd
Robert B. Schnabel

CU-CS-272-84



University of Colorado at Boulder
DEPARTMENT OF COMPUTER SCIENCE

* Research supported by NSF grant MCS 81-15475.
** Research supported by ARO contract DAAG 29-81-K-0108.

CONTINUITY OF THE NULL SPACE BASIS AND CONSTRAINED OPTIMIZATION

Richard H. Byrd*
Robert B. Schnabel**

Department of Computer Science
University of Colorado
Boulder, Colorado 80309

CU-CS-272-84 August, 1984

*Research supported by NSF grant MCS 81-15475
**Research supported by ARO contract DAAG 29-81-K-0108

WE ACKNOWLEDGE U. S. ARMY RESEARCH SUPPORT UNDER CONTRACT NO. DAAG 29-81-K-0108.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
Continuity of the samll space basis and constrained optimization

5. TYPE OF REPORT & PERIOD COVERED

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
Richard H. Byrd, Robert B. Schnabel

8. CONTRACT OR GRANT NUMBER(s)
DAAG-84-K-0140

9. PERFORMING ORGANIZATION NAME AND ADDRESS
University of Colorado, Computer Science
Campus Box 430, Boulder CO 80309

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
U. S. Army Research Office
Post Office Box 12211
Research Triangle Park, NC 27709

12. REPORT DATE
Sept. 84

13. NUMBER OF PAGES
14

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
NA

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Constrained optimization, null space basis, reduced Hessian

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
Many constrained optimization algorithms use a basis for the null space of the matrix of constraint gradients. Recently, methods have been proposed that enable this null space basis to vary contrinuously as a function of the iterates in a neighborhood of the solution. This paper reports results from topology showing that, in general, there is no continuous function that generates the null space basis of all full rank rectangular matrices of a fixed size. Thus constrained optimization algorithms cannot assume an everywhere continuous null space basis. We also give some indication of where these discontinuities must

occur. We then propse an alternative implementation of a class of constrained optimization algorithms that uses approzimations to the reduced Hessian of the Lagrangian but is independent of the choice of null space basis. This approach obviates the need for a continuously varying null space basis.

Abstract

Many constrained optimization algorithms use a basis for the null space of the matrix of constraint gradients. Recently, methods have been proposed that enable this null space basis to vary continuously as a function of the iterates in a neighborhood of the solution. This paper reports results from topology showing that, in general, there is no continuous function that generates the null space basis of all full rank rectangular matrices of a fixed size. Thus constrained optimization algorithms cannot assume an everywhere continuous null space basis. We also give some indication of where these discontinuities must occur. We then propose an alternative implementation of a class of constrained optimization algorithms that uses approximations to the reduced Hessian of the Lagrangian but is independent of the choice of null space basis. This approach obviates the need for a continuously varying null space basis.

1. Introduction.

Many recent papers on optimization algorithms use a basis for the null space of a given matrix. In these contexts the matrix often is the transpose of an $n \times t$ matrix $A(x)$ whose columns are the gradients of nonlinear constraints at a point x. The required basis for the null space of $A(x)^T$ forms an $n \times (n-t)$ matrix Z such that $A(x)^T Z = 0$. When x and thus A changes it is desirable to have the matrix Z change with it in a smooth manner. Indeed, in analysis of optimization methods, several authors (Coleman and Conn (1984), McCormick (1980), Nocedal and Overton (1983), Womersley and Fletcher (1982)) assume existence locally of a continuous function $Z(x)$ that gives the null space basis. Thus an important question is:

> Question 1 : Given $n > t \geq 1$, is it possible to compute the basis $z(A)$ for the null space of a matrix A^T as a continuous function defined for all full rank values of $A \in R^{n \times t}$?

(It should be pointed out that the null space itself and the orthogonal projection onto it are uniquely defined; a basis for the null space is not uniquely defined and so the possibility of a continuously varying choice of this basis is in question.)

This question was addressed in a recent paper by Coleman and Sorenson (1984). They point out that in several standard methods for computing Z (usually as a by-product of the QR factorization of A), Z is not a continuous function of A. However they also show that given any full rank matrix $\hat{A}$, it is possible to compute Z as a continuous function of A for all A in some neighborhood of $\hat{A}$, although this may not be of practical value in cases when $\hat{A}$ depends upon the solution of the problem being solved. Gill, Murray, Saunders, and Wright (1983) suggest a method which, given a sequence $\{A_k\}$,

generates Z_k as a function of A_k and the QR factorization of A_{k-1}. With this scheme, they show that if $A_k \to A^*$, and $\sum_{k=1}^{\infty} ||A_k - A^*|| \leq \infty$, then the sequence $\{Z_k\}$ converges. This last method appears to be a practical, if perhaps expensive, approach. Note that that with this approach, Z is no longer a function of A and it is possible that for distinct indices k and l, $A_k = A_l$ (or even $x_k = x_l$) but $Z_k \neq Z_l$. Thus the question of whether Z can be defined as a continuous function of A remains to be answered.

In the next section, we first report that several results from topology show that the answer to Question 1 is "no". This answer also implies that the QR factorization of a matrix cannot be given by a function that is continuous everywhere. We also give a proof of part of this result that shows, in addition, that the discontinuities of $z(A)$ can be numerous and widespread.

Section 3 discusses the dependence of some constrained optimization algorithms on the null space basis. Successive quadratic programming algorithms that use the reduced Hessian of the Lagrangian are independent of the choice of null space basis, but similar algorithms that use a quasi-Newton approximation to the reduced Hessian depend strongly upon the choice of this basis. We propose a modification to these methods which makes the iterates independent of the choice of the null space basis. This approach obviates the need for a continuously varying null space basis.

2. The possibility of a continuous null space basis.

Here we consider a function that gives a basis for all or a part of the null space of a matrix, and look at the question of what continuity properties it can have. Let z be a matrix function from $R^{n \times t}$ to $R^{n \times s}$, where $n > t$ and $s \leq n - t$, with the following properties.

Properties of a partial null space basis function.

1. $z(A) \in R^{n \times s}$ is defined for all full rank matrices in $R^{n \times t}$.
2. $z(A)$ has full rank s for all such matrices A.
3. $A^T z(A) = 0$ for all A in the domain of z.

If $s = n - t$ we will refer to Z as a null space basis function.

The question of whether such a function $z(A)$ can be continuous is an important question in topology, and was studied by many mathematicians in this century. The full answer, found by combining results in Eckmann (1943), Adams (1962), and Whitehead (1963), is stated below.

Theorem 2.1 Suppose z is a function from $R^{n \times t}$ to $R^{n \times s}$ satisfying properties 1-3, where $n \geq 2$, $1 \leq t < n$, $1 \leq s \leq n - t$. A continuous function z satisfying these properties exists in the following cases, and no others.

$n \geq 2, \quad t = n-1, \quad s = n-t = 1,$

$n = 7, \quad t = 2, \quad s = 1,$

$n = 8, \quad t = 3, \quad s = 1,$

$n = (2a+1)2^b 16^c, \quad t = 1, \quad s \leq 2^b + 8c - 1,$

where $a \geq 0$, $b = 1$, 2, or 3, $c \geq 0$.

An immediate corollary of Theorem 2.1 is that the answer to Question 1 is "yes" in the cases when the null space has dimension one or when A is 4×1 or 8×1, and "no" in all other cases. Thus in general, one may not assume that $z(A)$ is continuous. (When $t = n-1$, a continuous $z(A)$ can be constructed by a variant of the method of cofactors for calculating the inverse of a matrix. In the 4×1 case the null space basis is given by.

$$z(A) = \begin{bmatrix} a_2 & a_3 & a_4 \\ -a_1 & -a_4 & a_3 \\ a_4 & -a_1 & -a_2 \\ -a_3 & a_2 & -a_1 \end{bmatrix}$$

where a_1, a_2, a_3, and a_4 are the components of A. This form is actually

related to quaternions; the null space basis in the 8×1 case has a similar form and is related to Cayley numbers.)

Since the last $n-t$ columns in the orthogonal matrix of a QR factorization of an $n \times t$ matrix A provides a null space basis, it follows from the above theorem that the QR factorization of A with $n > t+1$ cannot, in general, be given by an everywhere continuous function of A. However, the triangular factor and the first t columns of the orthogonal factor can be computed by a function which is continuous at all full rank A. For example, Gram-Schmidt orthogonalization of the columns of A has this property.

Theorem 2.2 proves a stronger version of Theorem 2.1 in the cases when $n-t$ is even. We show that any function satisfying properties 1-3 has many discontinuities in this case. These discontinuities must occur within a specified distance of any given matrix, and they may occur at quite well conditioned matrices. The norms referred to are the Euclidean vector norm and the corresponding induced matrix norm.

Theorem 2.2 Suppose z is a function from $R^{n \times t}$ to $R^{n \times s}$ satisfying properties 1-3, with $n-t$ even, and a given matrix $\bar{A} \in R^{n \times t}$ has smallest singular value $\sigma_t > 0$. Then for any $\alpha > 0$, z has a discontinuity at some matrix within the set $\{A: \|A-\bar{A}\| \leq \sigma_t + \alpha\}$, and the discontinuity occurs at a matrix with smallest singular value equal to $\min\{\sigma_{t-1}, \alpha\}$.

Proof. Express the singular value decomposition of $\bar{A}$ by

$$\bar{A} = U \Sigma V^T,$$

where U and V are orthogonal with dimension $n \times t$ and $t \times t$ respectively and Σ is a $t \times t$ diagonal matrix with diagonal elements $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_t > 0$. Express U in terms of its columns as $U = [u^1\ u^2\ \cdots\ u^t]$. Now consider the matrix function $\tilde{A}(u)$ defined by

$$\tilde{A}(u) = [u^1\ u^2\ \cdots\ u^{t-1}\ u]\ diag(\sigma_1,\sigma_2,\ldots,\sigma_{t-1},\alpha)\ V^T. \qquad (2.1)$$

Let S denote the $n-t+1$ dimensional subspace $\{u\in R^n: u^T u^i = 0,\ i=1,\ \ldots,\ t-1\}$. Now define the function F on S by

$$F(u) = \frac{z(\tilde{A}(u))^1}{||z(\tilde{A}(u))^1||}$$

where the superscript 1 again denotes the first column. Note that since $z(A)$ has full rank the denominator is never zero, so $||F(u)|| = 1$ for all $u\in S$. Note also that, since the columns of $z(A)$ are orthogonal to the columns of A for any A, and thus to the columns of U, $F(u)\in S$ for all $u\in S$.

Thus the function F maps the unit sphere in S into the unit sphere in S. According to a well known theorem in topology (see for example Hilton and Wylie (1962)) any map of the sphere in an odd dimensional space into itself such that $F(u)$ is orthogonal to u for all u in the sphere cannot be continuous. Thus, if $n-t$ is even, any function $z(A)$ satisfying Properties 1-3 will have a discontinuity at $\tilde{A}(u)$ for some choice of u with $||u|| = 1$. Now note that $||\tilde{A}(u)-\bar{A}||\leq\sigma_t+\alpha$ and that the smallest singular value of $\tilde{A}(u)$ is $\min\{\sigma_{t-1},\ \alpha\}$. ■

This result is fairly sharp in that it shows that $z(A)$ must have a discontinuity in any ball around $\bar{A}$ of radius greater than σ_t, while it is possible to construct a continuous $z(A)$ on any ball around $\bar{A}$ with radius less than σ_t. This fact follows from a remark of Eckmann (1942). The following theorem shows the construction of such a function.

Theorem 2.3. Let $\bar{A}\in R^{n\times t}$ have smallest singular value $\sigma_t > 0$, and let $\bar{Z}$ be any null space basis for $\bar{A}$. Then the function

$$z(A) = [I-A(A^TA)^{-1}A^T]\bar{Z}, \qquad (2.2)$$

is a continuous null space basis function over all A such that $||A-\bar{A}|| < \sigma_t$.

Proof. Obviously, $A^T z(A) = 0$ for all A. To show where $z(A)$ has full rank, suppose that $z(A)u = 0$ for some A and some nonzero $u \in R^t$. Since the null space of $I - A(A^TA)^{-1}A^T$ is equal to the range of A then $\bar{Z}u = Ay$ for some $y \in R^n$. This implies that $\bar{A}^T Ay = \bar{A}^T \bar{Z}u = 0$. But then

$$\begin{aligned} ||A-\bar{A}||^2 &\geq \frac{||(A-\bar{A})y||^2}{y^Ty} \\ &= \frac{y^T(A^TA+\bar{A}^T\bar{A})y}{y^Ty} \\ &\geq \sigma_t^2. \end{aligned}$$

Thus $z(A)$ cannot be rank deficient if $||A-\bar{A}|| < \sigma_t$. ■

Note that the function $z(A)$ does not give an orthonormal basis for the null space of A, but to orthonormalize the basis by, say, Gram-Schmidt would preserve the continuity. It is interesting to note that the function $z(A)$ defined by equation (2.2) actually gives the null space basis for A which minimizes $||Z-\bar{Z}||$. (A similar construction of $z(A)$,

$$z(A) = [I-\bar{A}(A^T\bar{A})^{-1}A^T]\bar{Z},$$

which was used by Goodman (1982) in deriving a Newton method for constrained optimization, can also be shown to be a continuous null space basis over the same region.)

While Theorems 2.1 and 2.2 conclusively answer Question 1, a relevant question for many theoretical results in constrained optimization is:

> Question 2 : Given a continuous function $A(x) : D \subseteq R^n \rightarrow R^{n \times t}$, is it possible to compute the basis $z(x)$ for the null space of $A(x)^T$ as a continuous function of $x \in D$?

Obviously this is a different question than Question 1. There are clearly some functions $A(x)$ for which a continuous $z(x)$ exists (e.g. $A(x)$ = a constant matrix or a diagonal matrix function of x). Indeed Wajewski (1935) shows that if D is homeomorphic to a sphere, and if $A(x)$ has full rank for all $x \in D$,

then such a function exists. However the general answer to Question 2 also is "no". For example, using the same notation as in the proof of Theorem 2.2, for a given $\bar{A}$, consider a matrix function of the form $A(x) = \tilde{A}(x)$ as defined by (2.1). Then the function given by the first column of $z(x)$ maps R^n into the unit sphere in R^n. Now if, as before, we restrict x to lie in the unit sphere in the the subspace S we have a map of the unit sphere in that space into itself which must have a discontinuity if the dimension of the null space, $n-t$, is odd. Thus the assumption that the function $z(x)$ is continuous is not justified in general.

3. The dependence of some optimization algorithms on the null space basis.

One of the main motivations for wanting a smoothly varying null space basis arises in nonlinearly constrained optimization algorithms that make use of an approximation to the reduced Hessian of the Lagrangian. Consider the problem

$$\begin{aligned} &\underset{x \in R^n}{\text{minimize}} \quad f(x) \\ &\text{subject to} \quad c(x) = 0. \end{aligned} \tag{3.1}$$

where f is a real-valued function on R^n and c maps R^n to R^t, where $t < n$. Let $A(x)$ denote the $n \times t$ matrix whose columns are the gradients of the constraints. Given an estimate x_k to the solution of (3.1), let Z_k be a basis for the null space of $A(x_k)^T$, that is an $n \times (n-t)$ full rank matrix such that $A(x_k)^T Z_k = 0$. Then the reduced Hessian of the Lagrangian for problem (3.1) at x_k relative to the null space of $A(x_k)^T$ may be expressed as $Z_k^T \nabla^2 L(x_k, \lambda_k) Z_k$. Note that the reduced Hessian is dependent on the choice of null space basis. We will assume in this section that Z_k is chosen to have orthonormal columns.

Algorithms for solving (3.1) which make use of approximations to a reduced Hessian of the Lagrangian have been proposed by Coleman and Conn (1984), by Womersley and Fletcher (1982) and by Nocedal and Overton (1983). These algorithms are either a special case of, or are very similar to successive quadratic programming, and have the following form.

Algorithm 1.

Given x_k, let

$$h_k = -Z_k M_k^{-1} Z_k{}^T \nabla f(x_k)$$
$$v_k = -A_k (A_k^T A_k)^{-1} c(x_k) \tag{3.2}$$

or

$$v_k = -A_k (A_k^T A_k)^{-1} c(x_k + h_k) \tag{3.3}$$

$$x_{k+1} = x_k + h_k + v_k .$$

Here, M_k is an approximation to $Z_k^T \nabla^2 L(x_k, \lambda_k) Z_k$. Note that h_k is a solution of the homogeneous equality constrained quadratic program

$$\underset{h \in R^n}{\text{minimize}} \ \nabla f(x_k)^T h + \tfrac{1}{2} h^T Z_k M_k Z_k{}^T h \tag{3.4}$$

$$\text{subject to } A_k^T h = 0.$$

Also note that if the "vertical step", v_k is determined by (3.2) as opposed to (3.3) this class of algorithms is a special case of successive quadratic programming.

We are interested in the dependence of the step $(x_{k+1} - x_k)$ on Z_k and on the continuity of our choice of Z_k. An important special case is when

$$M_k = Z_k{}^T B_k Z_k$$

where B_k is an $n \times n$ matrix which does not depend on Z_k. For example, B_k might equal $\nabla^2 L(x_k, \lambda_k)$ or the identity matrix. In this case h_k is the solution to

$$\underset{h \in R^n}{\text{minimize}} \ \nabla f(x_k)^T h + \tfrac{1}{2} h^T B_k h$$

$$\text{subject to } A_k^T h = 0,$$

and thus, in exact arithmetic, h_k and the entire algorithm is independent of the choice of null space basis Z_k.

However, when one tries to generate M_k by quasi-Newton updates the step tends to depend strongly on the choice of Z_k. Suppose the reduced Hessian approximation M_k is updated by a method such as DFP or BFGS so that M_{k+1} satisfies

$$M_{k+1} Z^T s_k = Z^T y_k,$$

where Z is either Z_k or Z_{k+1}. Reasonable choices for s_k and y_k are

$$s_k = (x_{k+1} - x_k) \tag{3.5}$$

and

$$y_k = \nabla L(x_{k+1}, \lambda_k) - \nabla L(x_{k+1} - h_k, \lambda_k). \tag{3.6}$$

Other choices are suggested by Conn and Coleman (1984), by Womersley and Fletcher (1982), and, comprehensively, by Nocedal and Overton (1983). Now if M_k is a good approximation to $Z_k^T \nabla^2 L(x_k, \lambda_k) Z_k$, and if Z_{k+1} is very different from Z_k, even with a rank two update we cannot expect M_{k+1} to give a good approximation to ${Z_{k+1}}^T \nabla^2 L(x_k, \lambda_k) Z_{k+1}$.

Clearly it would be preferable for a reduced Hessian based algorithm not to depend on the choice of the null space basis at all. Such an approach is possible if, at each step, we regard the matrix M_k as implicitly giving an approximation B_k to the entire Hessian. Note that if M_k is an approximation to $Z_k^T \nabla^2 L(x_k, \lambda_k) Z_k$ then $Z_k M_k {Z_k}^T$ is an approximation to $P_k^T \nabla^2 L(x_k, \lambda_k) P_k$, where

$$P_k = Z_k {Z_k}^T = I - A(x_k)[A(x_k)^T A(x_k)]^{-1} A(x_k)^T$$

is the orthogonal projector onto the null space of $A(x_k)^T$, and is independent of the choice of Z_k.

Therefore we define

$$P_k B_k P_k = Z_k M_k Z_k{}^T$$

and extend our definition of B_k to approximate the rest of the Hessian. The idea of a reduced Hessian method is to ignore the contribution of $P_k \nabla^2 L(x_k,\lambda_k)(I-P_k)$ to the Hessian so we will take

$$P_k B_k (I-P_k) = (I-P_k) B_k P_k = 0.$$

Now the term $(I-P_k)B_k(I-P_k)$ has no effect on the current step, but it could affect all parts of B_{k+1} and thus subsequent steps. Just as a scaled identity is often used as an initial Hessian approximation, we will let

$$(I-P_k)B_k(I-P_k) = \beta_k(I-P_k)I(I-P_k) = \beta_k(I-P_k).$$

The scaling factor β_k is an approximation to $||\nabla^2 L(x_k,\lambda_k)||$; for example $\beta_k = ||M_k||$ is reasonable and minimizes the condition number of $||B_k||$. Thus, given M_k, our implicit Hessian approximation is

$$\begin{aligned} B_k &= P_k B_k P_k + P_k B_k (I-P_k) + (I-P_k) B_k P_k + (I-P_k) B_k (I-P_k) \\ &= Z_k M_k Z_k{}^T + \beta_k (I - Z_k Z_k{}^T). \end{aligned}$$

Now given the approximation B_k to the entire Hessian, the reduced Hessian approximation corresponding to a new basis Z_{k+1} is given by

$$Z_{k+1}{}^T B_k Z_{k+1} = T_k^T M_k T_k + \beta_k (I - T_k^T T_k)$$

where

$$T_k = Z_k{}^T Z_{k+1}.$$

This matrix may then be updated to give M_{k+1}.

The above logic leads to the following algorithm.

Algorithm 2

1. Given x_k and M_k compute h_k and v_k as in Algorithm 1.
2. Let $x_{k+1} = x_k + h_k + v_k$.
3. Compute Z_{k+1}, $T_k = Z_k{}^T Z_{k+1}$ and β_k.
4. Let

$$\bar{M}_k = T_k^T(M_k - \beta_k I)T_k + \beta_k I.$$

5. Update $\bar{M}_k$ to get M_{k+1} using the DFP or BFGS update,

with secant equation $M_{k+1}Z_{k+1}{}^T s_k = Z_{k+1}{}^T y_k$.

6. Set k to $k+1$ and go to step 1.

In step 5 a choice of s_k and y_k such as that given in (3.5) and (3.6) is reasonable as are any of the Z_{k+1} based choices given in Nocedal and Overton (1983, p.28). However if, instead, the update involves $Z_k{}^T s_k$ and $Z_k{}^T y_k$, as does the one used by used by Coleman and Conn (1984, p. 760), then the update in step 5 should come before step 3.

It should be noted that this algorithm involves three extra matrix multiplications over Algorithm 1 for a total of $2(n-t)^3+n(n-t)^2$ extra multiplications. If $n-t$ is small relative to n this is much less than the work involved in the computation of Z; it is also less work than the modification proposed by Gill, Murray, Saunders, and Wright to make Z vary smoothly, which was mentioned in Section 1. The benefit of this extra effort is that the algorithm is independent of the choice of basis as described below.

Theorem 3.1. If Algorithm 2 is used with the initial matrix $M_0 = Z_0{}^T B_0 Z_0$ where B_0 is independent of Z_0, and if s_k and y_k are given by (3.5)and (3.6), then the iterates produced are independent of the choice of Z_k for $k \geq 1$.

Proof. One need only note that Algorithm 2 is equivalent to the following algorithm if we let $B_k = Z_k M_k Z_k{}^T$.

Algorithm 3

1. Given x_k and B_k compute h_k by

$$\underset{h \in R^n}{\text{minimize}} \ \nabla f(x_k)^T h + \tfrac{1}{2} h^T B_k h$$

$$\text{subject to } A_k^T h = 0.$$

Compute v_k as in Algorithm 1.

2. Let $x_{k+1} = x_k + h_k + v_k$.

3. Compute P_{k+1}, and β_k.

4. Let

$$\bar{B}_k = P_{k+1}(B_k - \beta_k P_k)P_{k+1} + \beta_k I.$$

5. Update $\bar{B}_k$ using the DFP or BFGS update,

with secant equation $B_{k+1} P_{k+1} s_k = P_{k+1} y_k$.

6. Set k to $k+1$ and go to 1.

Note that this algorithm does not involve Z_k at all; thus the equivalent Algorithm 2 is independent of Z_k. ■

Algorithm 2 is very similar to the algorithms proposed by Coleman and Conn (1984), by Womersley and Fletcher (1982), and by Nocedal and Overton (1983). We have of course only described it very generally, but the main difference is the subspace shift in Step 4. Because of this it cannot be considered a special case of these algorithms, and questions of convergence need to be considered. We believe that the subspace shift is a good alternative to trying to ensure some kind of continuous change in the null space basis.

References

J. F. Adams (1962). "Vector Fields on Spheres", *Ann. of Math. 75*, pp. 603-632.

T.F. Coleman and A.R. Conn (1984). "On the local convergence of a quasi-Newton method for a nonlinear programming problem", *SIAM J. on Num. Anal. 21*, pp. 755-769.

T.F. Coleman and D.C. Sorenson (1984). "A note on the computation of an orthonormal basis for the null space of a matrix", *Mathematical Programming 29*, pp. 234-242.

B. Eckmann (1942). "Stetige Losungen linearer Gleichungssysteme", *Comment. Math. Helv. 15*, pp. 318-339.

P.E. Gill, W. Murray, M.A. Saunders, and M.H. Wright (1983). "On the representation of a basis for the null space," Report SOL 83-19, Stanford University.

J. Goodman (1983), "Newton's method for constrained optimization", Courant Institute of Mathematical Sciences, New York.

P.J. Hilton and S. Wylie (1962). *Homology Theory*, Cambridge University Press.

G. P. McCormick, (1980). "Locating an isolated global minimizer of a constrained nonconvex program", *Math. of Oper. Res. 5*, pp. 435-443.

J. Nocedal and M. Overton (1983). "Projected Hessian updating algorithms for nonlinearly constrained optimization", Report No.95, Computer Science Department, New York University.

T. Wajewski, (1935). "Sur les matrices dont les elements sont des fonctions

continues, *Comp. Math. II (1935)*, pp. 63-68.

G. W. Whitehead, (1963). "Note on cross-sections in Stiefel manifolds", *Comment. Math. Helv., 37, 239-240.*

R.S. Womersley and R. Fletcher (1982). "An algorithm for composite nonsmooth optimization problems", Dept. of Mathematics Report NA/60, University of Dundee.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE *(and Subtitle)*

Continuity of the null space basis and constrained optimization

5. TYPE OF REPORT & PERIOD COVERED

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)

Richard H. Byrd, Robert B. Schnabel

8. CONTRACT OR GRANT NUMBER(s)

DAAG-84-K-0140

9. PERFORMING ORGANIZATION NAME AND ADDRESS

University of Colorado, Computer Science
Campus Box 430, Boulder Co 80309

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS

U. S. Army Research Office
Post Office Box 12211
Research Triangle Park, NC 27709

12. REPORT DATE

Sept 85

13. NUMBER OF PAGES

13

14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*

15. SECURITY CLASS. *(of this report)*

Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

NA

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

NA

18. SUPPLEMENTARY NOTES

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Constrained optimization, null space basis, reduced Hessian

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

(see paper)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*